


## RESEARCH ARTICLE

# Nasopharyngeal Epstein-Barr virus DNA loads in high-risk nasopharyngeal carcinoma families: Familial aggregation and host heritability

Mei-Qi Zheng<sup>1,2</sup>  | Tong-Min Wang<sup>2</sup> | Ying Liao<sup>2</sup> | Wen-Qiong Xue<sup>2</sup> | Yong-Qiao He<sup>2</sup> | Zi-Yi Wu<sup>2</sup> | Da-Wei Yang<sup>2</sup> | Dan-Hua Li<sup>2</sup> | Chang-Mi Deng<sup>2</sup> | Yi-Jing Jia<sup>1,2</sup> | Lei-Lei Yuan<sup>1,2</sup> | Wen-Li Zhang<sup>2</sup> | Lu-Ting Luo<sup>1,2</sup> | Xia-Ting Tong<sup>1,2</sup> | Yan-Xia Wu<sup>2</sup> | Ting Zhou<sup>2</sup> | Xi-Zhao Li<sup>2</sup> | Ling-Ling Tang<sup>2</sup> | Jiang-Bo Zhang<sup>2</sup> | Yun-Fei Xia<sup>2,3</sup> | Jianbing Mu<sup>4</sup> | Wei-Hua Jia<sup>2</sup>

<sup>1</sup>School of Public Health, Sun Yat-sen University, Guangzhou, China

<sup>2</sup>State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>3</sup>Department of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>4</sup>Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland

## Correspondence

Wei-Hua Jia, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou 510060, China.  
Email: [jjawh@sysucc.org.cn](mailto:jjawh@sysucc.org.cn)

## Funding information

National Key Research and Development Program of China, Grant/Award Numbers: 2016YFC1302704, 2016YFC0902001; National Natural Science Foundation of China, Grant/Award Numbers: 81502056, 81802708, 81803319, 81903395, 81973131; Key Project for International Cooperation and Exchange of the National Natural Science Foundation of China, Grant/Award Number: 81220108022; Sino-Sweden Joint Research Programme, Grant/Award Number: 81861138006; Natural Science Foundation of Guangdong Province, Grant/Award Number: 2017A030312003; National Science Fund for Distinguished Young Scholars of China, Grant/Award Number: 81325018; Science and Technology Planning Project of Guangzhou, China, Grant/Award Number: 201804020094; Fundamental Research Funds for the Central Universities, Grant/Award Number: 19ykpy185

## Abstract

Nasopharyngeal carcinoma (NPC), the most common head and neck cancer, is characterized by distinct geographic distribution and familial aggregation. Multiple risk factors, including host genetics, environmental factor, and EBV infection, have been linked to the development of NPC, particularly in the familial clustering cases. However, the cause of NPC endemicity remains enigmatic due possibly to the complicated interplay between these risk factors. Recently, positive Epstein-Barr virus (EBV) DNA loads at nasopharyngeal (NP) cavity has been found to reflect NPC development and applied in NPC screening. To examine whether the increased NP EBV loads could aggregate in the families and be affected by host genetics and environmental factor, EBV loads were obtained by 510 NP brushing samples from eligible unaffected individuals, who have two or more relatives affected with NPC, in 116 high-risk NPC families. The correlation of relative pairs was estimated using S.A.G.E. (version 6.4, 2016), and host heritability of NP EBV loads was calculated with variance component models using SOLAR (version 8.4.2, 2019). In result, significant correlations of EBV loads were observed between parent-offspring pairs and sibling-sibling pairs ( $P < .001$ ), but not in distant kin relationship pairs. Interestingly, after excluding the shared environmental factor within families, host genetics contributes significantly to NP EBV loads with a heritability of 56.41% ( $P = 1.00 \times 10^{-7}$ ), and its effect was slightly

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Journal of Medical Virology* published by Wiley Periodicals LLC

elevated (68.86%,  $P = 3.40 \times 10^{-6}$ ) in families with more NPC cases ( $\geq 3$ ). These findings indicate that additional host-genetic variants involved in the EBV local NP mucosal behavior may be especially important for the development of NPC.

#### KEYWORDS

Epstein-Barr virus, familial aggregation, genetics, heritability, high-risk family, nasopharyngeal carcinoma

## 1 | INTRODUCTION

Nasopharyngeal carcinoma (NPC), a rare disease worldwide, has remarkably skewed geographic distribution and exhibits familial clustering in the endemic area.<sup>1,2</sup> Approximately 10% of cases have a positive family history, and individuals with a family history of NPC in first-degree relatives have a more than four-fold higher NPC risk compared with the general population.<sup>3</sup> It is suggested that host genetics and/or common environmental factors shared by the family members may contribute to the aggregation of NPC. Although a great effort has been made to clarify the genetic component, unfortunately, the identified risk alleles could only partially explain the heritability of NPC.<sup>4,5</sup> Unlike the functional mutations in the *BRCA* genes responsible for most inherited cases of breast and ovarian cancer,<sup>6,7</sup> the genetic mutations with large effects on familial NPC susceptibility remain unknown. Nevertheless, some clues were given by prior findings that a remarkable linkage signal of NPC in human leukocyte antigens (HLA) region was detected in affected sib-pairs from high-risk families.<sup>8</sup> Given the central roles of *HLA* genes in the body's immune response to pathogen attack,<sup>9</sup> the genetic risk variants may predispose individuals to NPC by modulating pathogen infections.

Epstein-Barr virus (EBV) is a member of the herpes virus family. As an oncogenic pathogen associated strongly with NPC, EBV, most infected latently in dysplastic epithelial cells, reactivates at mucosal surface parallelly with NPC formation.<sup>10,11</sup> The levels of viral biomarkers elevated before NPC onset, which could be used for early diagnosis.<sup>12,13</sup> Nasopharyngeal (NP) brushings, as a practical and minimally invasive approach for NPC diagnosis, showed the similar viral expression profile with the corresponding NPC biopsies, and EBV DNA loads in NP brushings were also similarly high with those in NPC biopsies.<sup>10</sup> We and others identified EBV DNA loads in NP brushings as an efficient NPC diagnostic biomarker with both high sensitivity and specificity.<sup>14,15</sup>

Given the close association between NP EBV loads and NPC, we reason that the elevated NP EBV loads may also show familial aggregation and be affected by the host genetics, among the individuals from high-risk families with multiple NPC cases, which may be associated with NPC familial aggregation.

To examine this hypothesis, we established an NPC family cohort in southern China. The familial aggregation and the host heritability of NP EBV loads was estimated among unaffected relatives from high-risk NPC families with two or more NPC cases. We find that NP EBV loads are positively correlated between relative pairs from NPC

susceptible families and have a significantly high heritability. We believe that these findings not only provide clues in the search of novel susceptibility genes in NPC patients with high-risk family history but, more importantly, they may improve our understanding of the genetic mechanisms for more widely distributed EBV-relevant diseases.

## 2 | MATERIALS AND METHODS

### 2.1 | High-risk NPC family screening program

In September 2016, we launched an NPC familial study in southern China, entitled "High-risk Nasopharyngeal Carcinoma Family Screening Program", to recruit NPC multiplex families with two or more NPC cases in their family in Sun Yat-sen University Cancer Center (SYSUCC, Guangzhou, China). We recruited patients with NPC as well as the unaffected relatives from high-risk families via handing out of brochures in our cancer center, posting our project information on the websites and our official social media (WeChat) account. People who registered for the program were telephone-reviewed to confirm their NPC family history, and those families with two or more affected members who have medical records with pathological diagnosis of NPC were enrolled.

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional research committee (the human ethics committee of the Sun Yat-sen University Cancer Center, GZR2012-008). This article does not contain any studies with animals performed by any of the authors.

All subjects were fully informed with a signed consent and accomplished a face-to-face interview by trained investigators using a structured questionnaire. This questionnaire contained the enrolled individuals' demographics, clinical history of ear-nose-throat (ENT) and oral disease, family history of malignancy, occupational history, and lifestyle information, including smoking status and alcohol consumption. The questionnaire also included name, gender, date of birth, and other general characteristics of all first-degree relatives of the enrolled individuals. Subsequently, we drew a pedigree structure for each enrolled family and determined the familial relationship among unaffected individuals based on the above information obtained from family members.

For NPC screening, we provided endoscopic examination of the nasopharynx, EBV antibody detection, and EBV copy number detection via NP brush samples. During the screening program, we

collected biological samples from available patients with NPC and their healthy relatives which comprised of 8 mL blood and NP swab/brushing swept more than five times over the surface of the pharyngeal recess under nasopharyngoscopy.

By October 2018, we recruited a total of 116 high-risk families which contain two or more NPC cases for each family. The eligible unaffected individuals from the enrolled families were used for the followed analyses.

The visualized enrollment of families and participants is shown in Figure 1.

## 2.2 | EBV DNA loads and VCA-IgA levels detections

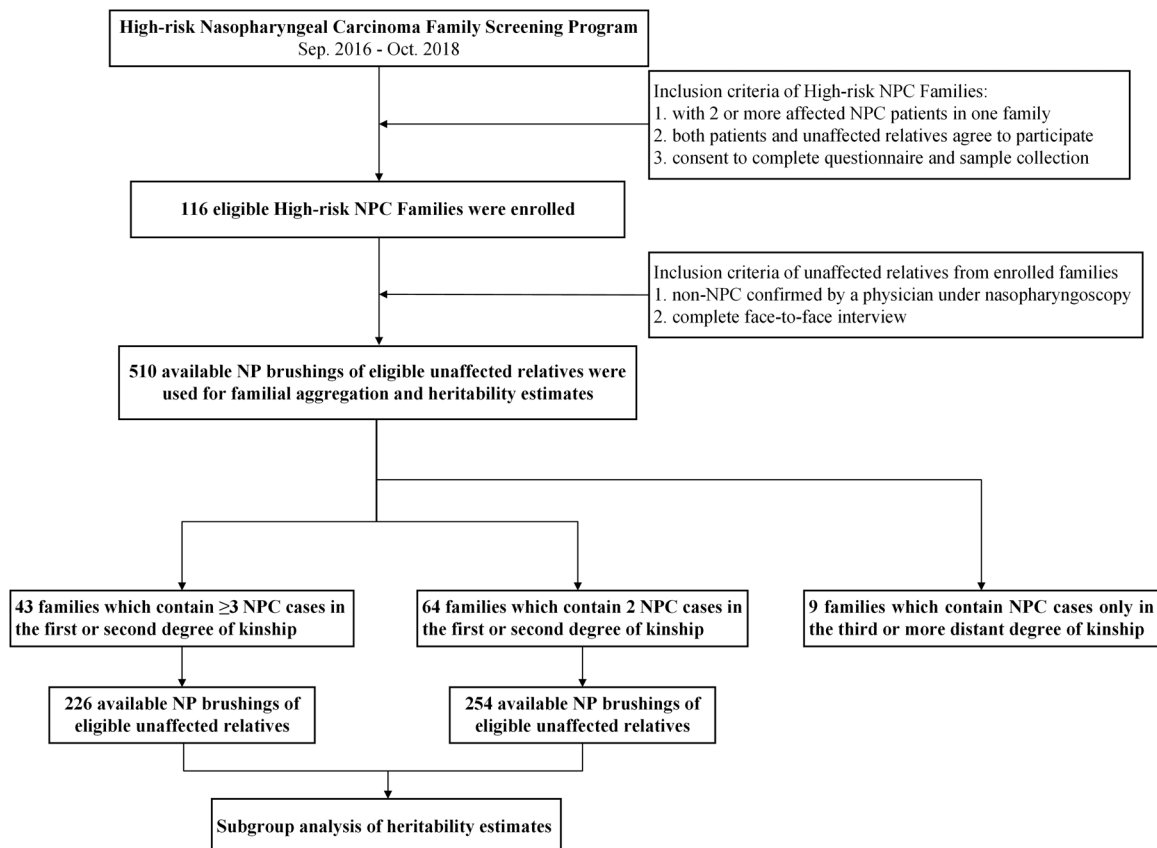
NP brushing DNA was extracted by an automated workstation (Chemagic Star; Hamilton Robotic, Bonaduz, GR, Switzerland) using the protocol recommended by the manufacturer. EBV DNA loads were detected by real-time quantitative polymerase chain reaction (qPCR) as described in the previous studies.<sup>16-18</sup> In brief, qPCR was performed for amplification targeting the BamHI-W region of EBV genome and a dual-labeled fluorescence probe was used for hybridization. The qPCR assay was conducted in 384-well plates including negative control and all samples tested in duplicate. We converted cycle threshold (CT) value of samples to DNA copy number using the established standard curve upon a series of standard samples for gradient EBV DNA copies

( $10^2$ ,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ , and  $10^7$  copies per  $2 \mu\text{L}$ ). After using the DNA concentration to standardize viral copy numbers, the EBV DNA loads were expressed as copy/ng DNA. Serum immunoglobulin A antibodies against EBV viral capsid antigen (VCA-IgA) levels were detected using an Anti-EBV-VCA ELISA (IgA) kit (EUROIMMUN Medizinische Labordiagnostika AG). Detailed descriptions of the qPCR assay are shown in the Supplementary Methods. The raw data in this article has been uploaded in the Research Data Deposit with an RDD number of RDDA2020001574.

## 2.3 | Statistical analysis

We classified the unaffected individuals into an EBV high-level group and low-level group according to the obtained median of EBV DNA loads in their NP brushings. The associations between NP EBV loads and demographics, anthropometries, chronic ENT disease status as well as EBV VCA-IgA levels were examined using a linear mixed model (LMM), which can handle the correlated family members by adjusting for the kinship structure among the subjects.<sup>19</sup>

Subsequently, we used the above unaffected individuals from the high-risk NPC families to construct relative pairs upon the familial relationship and estimate the familial correlation coefficients for NP EBV DNA loads and three chronic ENT disease occurrences (chronic rhinitis, allergic rhinitis, and chronic pharyngitis) between the relative



**FIGURE 1** The study design and the enrollment of unaffected individuals from high-risk NPC families. NPC, nasopharyngeal carcinoma

pairs by the FCOR program of Statistical Analysis for Genetic Epidemiology (S.A.G.E., version 6.4, 2016). The correlation coefficients were visualized by the "forestplot" package in R (version 3.5.3).

We further estimated the heritability of the above phenotypes, which is a measure of how well differences in people's genes account for differences in their phenotypes. A variance component model in Sequential Oligogenic Linkage Analysis Routines (SOLAR, version 8.4.2, 2019)<sup>20</sup> was used. SOLAR models the phenotypic covariance matrix as

$$\Omega = 2\Phi\sigma_g^2 + H\sigma_c^2 + I\sigma_e^2,$$

where  $\Omega$  denotes a phenotypic covariance matrix,  $\Phi$  is the expected kinship matrix derived from pedigree information,  $\sigma_g^2$  represents genetic variance,  $H$  is a matrix derived from cohabitation information among family members,  $\sigma_c^2$  represents variance due to shared environmental influences within families,  $I$  is an identity matrix, and  $\sigma_e^2$  represents variance due to unique environmental influences of measured covariates as well as of random/unmeasured variates.

SOLAR decomposes the phenotypic variance ( $\sigma_p^2$ ) using the additive assumption of the variance components ( $\sigma_p^2 = \sigma_g^2 + \sigma_c^2 + \sigma_e^2$ ) to calculate heritability with  $h^2 = \sigma_g^2/\sigma_p^2$  for NP EBV loads and chronic ENT diseases occurrences, which represents the proportion of phenotypic variance explained by the host-genetic effect. A household effect component ( $c^2$ ) reflecting the shared environmental exposure within families was calculated with  $c^2 = \sigma_c^2/\sigma_p^2$  simultaneously based on self-reported cohabitation. For age, sex, smoking, and drinking which may correlate with the investigated phenotypes, we put those into the initial variance component models as covariates. The models would examine the significance of each covariate effect, and the significant covariates were kept in the final models for heritability estimates and for the computation of the proportion of phenotypic variance explained by the significant covariate effect ( $e^2$ ). In addition, we selected EBV VCA-IgA, given its reported familial correlations,<sup>21</sup> to estimate its heritability and to adjust VCA-IgA levels as an additional covariate for heritability estimates of NP EBV loads. For height and body mass index (BMI) the heritability was estimated as well to verify the validity of these analyses given their reported high heritability and the multiple identified genetic susceptibility variants.<sup>22-24</sup>

At last, a subgroup analysis was conducted for the phenotypes with significant heritability. We classified families with two NPC cases and families with more than or equal to three NPC cases into two subgroups and then conducted heritability estimates in each group (Figure 1).

Detailed descriptions of the statistical analyses are shown in the Supplementary Methods.

### 3 | RESULTS

#### 3.1 | Characteristics of the study population

A total of 510 NP brushings and corresponding blood samples from the unaffected individuals of 116 high-risk NPC families were

collected. We constructed 384 first-degree relative pairs among the recruited individuals, including 154 parent-offspring pairs and 230 sibling-sibling pairs. Besides, 250 aunt/uncle-nephew/niece (avuncular) pairs and 375 cousin-cousin pairs were for more distant relative pairs. Overall, a total of 1136 relative pairs were constructed among 510 unaffected individuals whose NP brushing samples were eligible (Table 1).

EBV DNA was detectable in 387 NP brushing samples, with 75.88% (387/510) of the detection rate similar to the reported rates of EBV-DNA detectable NP brushing samples of non-NPC controls ranging from 70.6% to 77%.<sup>15,25</sup> The median of EBV DNA loads was 0.42 (ranging from 0 to 111.08) among unaffected individuals. There was a significant association between EBV DNA loads in NP brushings and serum VCA-IgA levels ( $P = .003$ ), but we did not find any significant associations between EBV DNA loads and individuals' characteristics nor chronic ENT diseases (Table 2).

#### 3.2 | Familial aggregation for NP EBV DNA loads

To identify the aggregation of NP EBV loads between the relatives from high-risk NPC pedigrees, we evaluated the correlation coefficients for parent-offspring, sibling-sibling, aunt/uncle-nephew/niece (avuncular), and cousin-cousin pairs. For NP EBV loads, significant first-degree relative correlations were found with the coefficients ( $r$ ) of 0.34 ( $P = .0007$ ) between parent-offspring pairs as well as 0.27 ( $P = .0002$ ) between sibling-sibling pairs. The positive correlation coefficient was degressive when the kin relationship became more distant, which were 0.17 for second-degree relative (avuncular) pairs and 0.12 for third-degree relative (cousin-cousin) pairs (Figure 2).

The familial correlations for other phenotypes did not change with the degree of relatedness between kin in pairs as those for NP EBV loads did. The occurrences of the analyzed chronic diseases

**TABLE 1** The numbers of examined relative pairs for nasopharyngeal brushings

Relationship <sup>a</sup>	No. of relative pairs <sup>b</sup>
1st-degree pairs	384
Parent-Offspring	154
Sibling-Sibling	230
2nd-degree pairs	274
Avuncular	250
3rd-degree pairs	399
Cousin-Cousin	375
Other relative pairs	79
Total relative pairs	1136

<sup>a</sup>The relative pair numbers were counted using SOLAR software according to kin's matrix.

<sup>b</sup>Nasopharyngeal brushings were collected from 510 healthy individuals from the high-risk NPC pedigrees.

**TABLE 2** Characteristics of EBV (DNA load) high-level group and low-level group

Characteristics <sup>a</sup>	EBV high-level (no.) (%) <sup>b</sup>	EBV low-level (no.) (%) <sup>b</sup>	OR (95% CI) <sup>c</sup>	P-value <sup>c</sup>
EBV DNA loads, median (IQR)	1.38 (0.71, 3.71)	0.01 (0, 0.21)		
EBV VCA-IgA levels, median (IQR)	0.66 (0.47, 0.94)	0.57 (0.43, 0.84)		
Negative	106 (41.73)	141 (55.51)	Ref.	Ref.
Positive	148 (58.27)	113 (44.49)	1.54 (1.16-2.04)	.003
Age range, y, median (IQR)	35 (27, 47)	35 (27, 46)		
<18	14 (5.51)	17 (6.67)	Ref.	Ref.
18-30	79 (31.10)	77 (30.20)	1.18 (0.50-2.81)	.708
31-45	89 (35.04)	95 (37.25)	1.14 (0.48-2.69)	.769
>45	72 (28.35)	66 (25.88)	1.25 (0.52-3.01)	.611
Sex				
Male	117 (45.88)	134 (52.55)	Ref.	Ref.
Female	138 (54.12)	121 (47.45)	1.24 (0.94-1.63)	.125
Smoking				
Never smoker	186 (73.52)	183 (72.62)	Ref.	Ref.
Current smoker	52 (20.55)	54 (21.43)	0.92 (0.56-1.50)	.734
Ex-smoker	15 (5.93)	15 (5.95)	0.95 (0.41-2.18)	.894
Drinking				
Never drinking	111 (43.70)	113 (44.31)	Ref.	Ref.
Regular drinking	49 (19.29)	49 (19.22)	0.93 (0.55-1.59)	.798
Occasional drinking	94 (37.01)	93 (36.47)	0.99 (0.64-1.53)	.968
Height, cm, median (IQR)	163 (158, 168)	165 (158, 170)	0.98 (0.96-1.01)	.167
BMI, kg/m <sup>2</sup> , median (IQR)	22.03 (20.00, 24.29)	22.10 (19.72, 24.24)	1.01 (0.95-1.08)	.674
Chronic rhinitis				
No	196 (83.76)	206 (86.19)	Ref.	Ref.
Yes	38 (16.24)	33 (13.81)	1.15 (0.66-2.00)	.616
Allergic rhinitis				
No	201 (85.90)	203 (84.94)	Ref.	Ref.
Yes	33 (14.10)	36 (15.06)	0.88 (0.50-1.56)	.671
Chronic pharyngitis				
No	187 (78.57)	196 (84.12)	Ref.	Ref.
Yes	51 (21.43)	37 (15.88)	1.47 (0.88-2.49)	.141

Abbreviations: BMI, body mass index; CI, confidence interval; EBV, Epstein-Barr virus; IQR, interquartile range; OR, odds ratio; Ref., reference.

<sup>a</sup>EBV DNA loads: Epstein-Barr virus DNA loads detected in nasopharyngeal brushings, expressed as copy/ng DNA; EBV VCA-IgA levels: serum levels of immunoglobulin A against viral capsid antigen of Epstein-Barr virus, expressed as a ratio of the sample absorbance to standard reference material absorbance; *Positive* and *Negative* were classified according to the obtained median of EBV VCA-IgA levels from the enrolled unaffected individuals; median and IQR were used to describe the distribution of continuous variables; *Current smoker*: the person who had ever smoked at least 60 cigarettes in 6 mo and stopped smoking for less than 1 y; *Ex-smoker*: the person who had stopped smoking for at least 1 y; *Never smoker*: the person who had never smoked more than 60 cigarettes in 6 mo; *Regular drinking*: the person who had ever drunk at least once a week in 6 mo; *Occasional drinking*: the person who had ever drunk less than once a week in 6 mo but more than 10 times per year; *Never drinking*: the person who had never drunk more than 10 times per year.

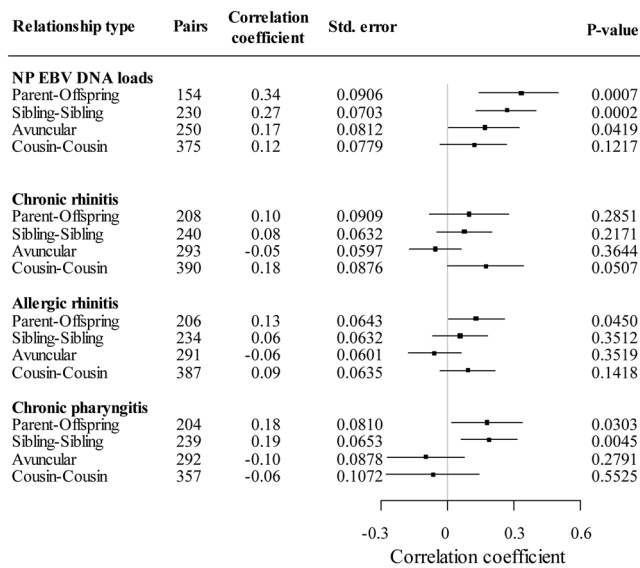
<sup>b</sup>EBV high-level represents a group of the unaffected individuals whose EBV DNA loads were higher than the obtained median from the enrolled unaffected individuals, others whose EBV DNA loads were lower than the obtained median belong to EBV low-level group.

<sup>c</sup>The associations between characteristics (explanatory variables) and NP EBV DNA levels (the binary response variable) were estimated by "lme4" package in R (version 3.5.3) using a generalized linear mixed model.

were weakly correlated for most types of relative pairs. It was only slightly significant between first-degree relative pairs for chronic pharyngitis (parent-offspring pairs:  $r = .18$ ,  $P = .0303$ ; sibling-sibling pairs:  $r = .19$ ,  $P = .0045$ ) and between parent-offspring pairs for allergic rhinitis ( $r = .13$ ,  $P = .0450$ ) (Figure 2).

### 3.3 | Host heritability for NP EBV DNA loads

To understand the possible causes of the remarkable familial aggregation of NP EBV loads, we further estimated heritability ( $h^2$ ) to distinguish the host-genetic effect contributing to the EBV load variation among



**FIGURE 2** Familial aggregation estimates for NP EBV loads and chronic disease occurrences. The familial correlation coefficients of Parent-Offspring, Sibling-Sibling, aunt/uncle-nephew/niece (Avuncular), and Cousin-Cousin pairs were estimated with standard error (SE) for EBV DNA loads in nasopharyngeal brushings (NP EBV DNA loads), chronic rhinitis status, allergic rhinitis status, and chronic pharyngitis status. The estimations and the statistical tests were conducted using the FCOR program in the software Statistical Analysis for Genetic Epidemiology (S.A.G.E., version 6.4, 2016). EBV, Epstein-Barr virus; NP, nasopharyngeal

individuals by variance component models. We found a significant high host heritability of NP EBV DNA loads ( $h^2 = 59.45\%$ ,  $P = 1.79 \times 10^{-10}$ ). After excluding the proportion explained by shared environmental exposures within families (4.72% of the phenotypic variance), the model still showed a strikingly high heritability of EBV DNA loads ( $h^2 = 56.41\%$ ,  $P = 1.00 \times 10^{-7}$ ) (Table 3). Besides, the robust and larger heritability of EBV DNA loads was observed with 59.68% after adjusting for VCA-IgA levels ( $P = 1.34 \times 10^{-8}$ ).

**TABLE 3** The estimates of host heritability on investigated phenotypes

Phenotypes <sup>a</sup>	Sample size (no.) <sup>b</sup>	$h^2$ (%) (P-value) <sup>c</sup>	$c^2$ (%) (P-value) <sup>d</sup>	$e^2$ (%) <sup>e</sup>	$h^2$ with $c^2$ (%) (P-value) <sup>f</sup>
EBV DNA loads	510	59.45 ( $1.79 \times 10^{-10}$ )	4.72 (.285)	0.81	56.41 ( $1.00 \times 10^{-7}$ )
EBV VCA-IgA levels	562	35.18 ( $7.96 \times 10^{-5}$ )	2.89 (.366)	7.74	33.14 ( $1.53 \times 10^{-3}$ )
Chronic rhinitis	556	14.32 (.266)	9.94 (.339)	3.55	9.00 (.366)
Allergic rhinitis	560	38.19 (.043)	13.47 (.230)	0.84	29.54 (.167)
Chronic pharyngitis	551	38.46 (.025)	30.28 (.048)	3.40	26.31 (.115)
Height	584	55.66 ( $1.69 \times 10^{-12}$ )	0	40.53	55.66 ( $1.69 \times 10^{-12}$ )
BMI	579	34.71 ( $2.46 \times 10^{-5}$ )	0	27.38	34.71 ( $2.46 \times 10^{-5}$ )

Abbreviations: BMI, body mass index; EBV, Epstein-Barr virus.

<sup>a</sup>EBV DNA loads: Epstein-Barr virus DNA loads detected in nasopharyngeal brushings; EBV VCA-IgA levels: serum levels of immunoglobulin A against viral capsid antigen of Epstein-Barr virus.

<sup>b</sup>The samples with missing information of significant covariates were excluded from the analyses.

<sup>c</sup> $h^2$ : "heritability", the proportion of the phenotypic variance explained by host-genetic effect.

<sup>d</sup> $c^2$  indicates the proportion of the phenotypic variance explained by shared environmental effect of cohabitation.

<sup>e</sup> $e^2$  indicates the proportion of the phenotypic variance explained by significant covariates, and this proportion would be removed from the total variance of the phenotype before heritability estimates.

<sup>f</sup> $h^2$  was estimated with counting  $c^2$  in variance component models, whether  $c^2$  was significant or not.

Although the heritability of allergic rhinitis and chronic pharyngitis was estimated as being 38.19% ( $P = .04$ ) and 38.46% ( $P = .03$ ), respectively, no statistical significance was found after accounting for the household effect of shared environmental exposures within families ( $h^2 = 29.54\%$ ,  $P = .17$  for allergic rhinitis and  $h^2 = 26.31\%$ ,  $P = .12$  for chronic pharyngitis), suggesting that genetic factors had smaller contributions to the occurrences of chronic ENT diseases. As expected, a significant host heritability was estimated as being 33.14% for EBV VCA-IgA levels ( $P = .0015$ ), 55.66% for height ( $P = 1.69 \times 10^{-12}$ ), and 34.71% for BMI ( $P = 2.46 \times 10^{-5}$ ) after excluding the household effect (Table 3). The results of covariate screenings before the above heritability estimates of phenotypes are shown in Table S1.

As the elevation of NP EBV loads is a high-risk factor for NPC development, we further investigated whether the host-genetic effect on NP EBV loads would be correlated with NPC development and whether higher genetic contributions to viral loads would be observed in families with more NPC cases. Interestingly, we observed larger genetic effects on NP EBV loads in families with more than or equal to three NPC cases (family numbers = 43) with  $h^2 = 68.86\%$  ( $P = 3.40 \times 10^{-6}$ ), which are higher than those estimated from families with two cases (family numbers = 64,  $h^2 = 47.06\%$ ,  $P = 4.96 \times 10^{-4}$ ). Furthermore, NP EBV loads showed the remarkably higher heritability than other analyzed phenotypes in more than or equal to three NPC case families. A similar trend was found for serum VCA-IgA levels but not for height and BMI (Table 4). All significant results of hypothesis tests remained the statistical significance after adjusting P-values for multiple testing by the FDR method.

## 4 | DISCUSSION

This is the first report regarding the familial aggregations of EBV DNA loads in NP brushing samples and the possible causes, in a high-risk NPC family cohort. NP brushings provide the new approaches for early NPC diagnosis with the minimally invasive procedure.



**TABLE 4** Subgroup analysis of host heritability on EBV DNA loads and other phenotypes

Phenotypes <sup>a</sup>	Families with ≥ 3 NPC patients				Families with 2 NPC patients				
	Sample size (no.) <sup>b</sup>	$h^2$ (%) (P-value) <sup>c</sup>	$h^2$ with $c^2$ (%) (P-value) <sup>d</sup>	$c^2$ (%) (P-value) <sup>d</sup>	Sample size (no.) <sup>b</sup>	$h^2$ (%) (P-value) <sup>c</sup>	$h^2$ with $c^2$ (%) (P-value) <sup>d</sup>	$c^2$ (%) (P-value) <sup>d</sup>	$e^2$ (%) <sup>e</sup>
EBV DNA loads	226	68.86 (3.40 × 10 <sup>-6</sup> )	68.86 (3.40 × 10 <sup>-6</sup> )	0	254	58.59 (2.70 × 10 <sup>-6</sup> )	47.06 (4.96 × 10 <sup>-4</sup> )	18.58 (.059)	0.50
EBV VCA-IgA levels	253	40.43 (2.28 × 10 <sup>-3</sup> )	37.06 (.021)	3.59 (.385)	282	25.62 (.011)	25.62 (.011)	0	10.34
Height	258	51.42 (4.34 × 10 <sup>-5</sup> )	51.42 (4.34 × 10 <sup>-5</sup> )	0	287	55.53 (9.00 × 10 <sup>-7</sup> )	55.53 (9.00 × 10 <sup>-7</sup> )	0	40.68
BMI	263	29.26 (6.34 × 10 <sup>-3</sup> )	28.73 (.014)	0.99 (.460)	287	38.57 (1.15 × 10 <sup>-3</sup> )	38.57 (1.15 × 10 <sup>-3</sup> )	0	29.92

Abbreviations: BMI, body mass index; EBV, Epstein-Barr virus.

<sup>a</sup>EBV DNA loads: Epstein-Barr virus DNA loads detected in nasopharyngeal brushings; EBV VCA-IgA levels: serum levels of immunoglobulin A against viral capsid antigen of Epstein-Barr virus.

<sup>b</sup>The samples with missing information of significant covariates were excluded from the analyses.

<sup>c</sup> $h^2$ : "heritability", the proportion of the phenotypic variance explained by host-genetic effect.

<sup>d</sup> $c^2$  indicates the proportion of the phenotypic variance explained by shared environmental effect of cohabitation;  $h^2$  was estimated with counting  $c^2$  in variance component models, whether  $c^2$  was significant or not.

<sup>e</sup> $e^2$  indicates the proportion of the phenotypic variance explained by significant covariates, and this proportion would be removed from the total variance of the phenotype before heritability estimates.

EBV DNA loads quantified in NP brushings were similar to parallel biopsies and identified as a true and efficient diagnostic biomarker. It could directly reflect the EBV genome content at the surface of the primary tumor formation site of NPC, and its elevation was considered to originate from tumor cell instead of newly produced virions.<sup>10,15,25,26</sup> Therefore, NP EBV loads are a promising biomarker strongly correlated to NPC risk. In addition, our study population had a high prevalence of NPC which may increase the likelihood of identifying the genetic factors affecting the NP EBV DNA loads and therefore as the predisposing factor to NPC. Overall, the High-risk NPC Family Screening Program provides the valuable opportunity to identify the genetic determinants that underlie NPC development.

In the current study, we found that individual EBV DNA loads in NP brushings were significantly correlated between relative pairs. It showed the significant and positive correlations between parent-offspring pairs and between sibling-sibling pairs, respectively. This may be related to the close relatives who have a higher risk of carrying the same genetic variants or share common life-styles. In addition, the more distant the kin relationship between the relative pairs was, the smaller the magnitude of the correlation coefficient for NP EBV loads between the pairs was (Figure 2), suggesting that the familial aggregation of NP EBV loads could be more likely to attribute to the host-genetic factors. On the contrary, the occurrences of chronic rhinitis, allergic rhinitis, and chronic pharyngitis had low familial aggregations compared to NP EBV loads, suggesting that the predispositions of these diseases were more likely to attribute to the individual exposure rather than shared factors by relatives.

We subsequently estimated heritability for several phenotypes with a range of 9.00% to 56.41% after accounting for the shared environmental effect within families based on self-reported cohabitation. Given that heritability indicates a contribution of genetic effects on phenotypes, the largest host heritability (56.41%) of EBV loads in NP brushings strongly suggested that host-genetic factors are important determinants for the viral genome content in the brushings sampled from the NP surface, confirming that the substantial genetic factors contribute to the familial aggregations of NP EBV loads. Interestingly, after adjusting for the associated anti-VCA IgA antibody levels (OR = 1.54, 95% confidence interval (CI): 1.16-2.04), a robust heritability of NP EBV loads remained with 59.68%. This suggested that the lytic antibody in serum, VCA-IgA, might respond to EBV loads in NP brushings, but the host-genetic effect on them was more likely to be explained by distinct genes. Other covariates such as age, sex, smoking, and drinking status (Table S1) showed a small component of NP EBV load variance (0.81%), confirming the lack of associations between NP EBV loads and individuals' characteristics (Table 2). Consistent with the weak familial aggregations, allergic rhinitis and chronic pharyngitis had lower heritability, however, the heritability of all chronic ENT diseases was not significant once accounting for the shared environmental effect within families. It was accordant to the previous reports that home factors shared by cohabitants could have a substantial influence on the occurrences of chronic ENT diseases.<sup>26-28</sup> For other phenotypes, the host heritability of VCA-IgA levels of 33.14% was in the same range of the reported heritability of VCA-IgG

titers (from 0.32 to 0.48),<sup>29</sup> and the significantly large heritability of height and BMI was also in accordance with the previous reports.<sup>21-23</sup> The consistent results from height and BMI further support the reliability of host heritability estimates for NP EBV DNA loads in this familial population.

The results of subgroup analyses showed that NP EBV DNA loads seemed to be affected more largely by host-genetic effects in families with more than or equal to three patients with NPC compared with those in families with two patients with NPC. Notably, no such trend between the two groups existed in anthropometric phenotypes, and the differences of the heritability between NP EBV loads and other analyzed phenotypes became greater in more than or equal to three NPC families. It further suggests that the large host-genetic effect on EBV DNA loads detected at the surface of NP mucosa was likely to contribute to the predisposition and familial aggregation of NPC as well.

In conclusion, we found significant host-genetic effects on EBV DNA loads in NP brushing samples through the analysis of the familial correlations and host heritability estimates in high-risk NPC families. Furthermore, a larger heritability of NP EBV loads was exhibited in high-risk families with more affected individuals, supporting the notion that the substantial genetic susceptibility of NP EBV DNA loads may relate to NPC risk. Considering the prior reports that HLA genes have a consistent and strong association with NPC risk,<sup>8,30,31</sup> we also conducted an association analysis between 11 previously Genome-Wide Association Studies-identified single nucleotide polymorphisms<sup>32,33</sup> and NP EBV loads but did not find any variants with significance (data not shown). The genetic risk variants/mutations contributing to the heritability of NP EBV DNA loads have not been identified, which could be the limitation of our current study. Hence, more research works are required to identify the susceptibility alleles of EBV infection among NPC predisposed individuals. Overall, our findings from high-risk NPC pedigrees reveal that host-genetic factors play an important role in maintaining EBV genome loads in the NP surface, and it may cause elevated risk of NPC, particularly for the familial NPC.

#### ACKNOWLEDGMENTS

We thank Yu-Ying Liu and Yan-Hong Li at Department of Cancer Prevention Center, Sun Yat-sen University Cancer Center, Guangzhou, China for their help with the sample collection. This study was funded by grants from Sino-Sweden Joint Research Programme (81861138006); National Natural Science Foundation of China (81973131, 81803319, 81903395, 81502056, and 81802708); the Science and Technology Planning Project of Guangzhou, China (201804020094); Natural Science Foundation of Guangdong Province (2017A030312003); the Fundamental Research Funds for the Central Universities (19ykpy185); the Key Project for International Cooperation and Exchange of the National Natural Science Foundation of China (81220108022); National Science Fund for Distinguished Young Scholars of China (81325018); and National Key Research and Development Program of China (2016YFC1302704, 2016YFC0902001).

#### CONFLICT OF INTERESTS

All the authors declare that there are no conflict of interests.

#### AUTHOR CONTRIBUTIONS

This study was conceived and designed by W-HJ. The study design and data analysis were performed by M-QZ and T-MW. Material preparation, participant recruitment, sample collection and detection, and data collection were performed by M-QZ, T-MW, YL, W-QX, Z-YW, D-WY, D-HL, C-MD, Y-JJ, L-LY, W-LZ, L-TL, X-TT, Y-XW, TZ, X-ZL, L-LT, Y-FX, Y-QH, and J-BZ. W-HJ, M-QZ, and T-MW wrote the article. JM gave critical revision of the article. All authors read and approved the final manuscript.

#### ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee (the human ethics committee of the Sun Yat-Sen University Cancer Center, GZR2012-008) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

#### ORCID

Mei-Qi Zheng  <http://orcid.org/0000-0002-7595-8120>

#### REFERENCES

1. Tang LL, Chen WQ, Xue WQ, et al. Global trends in incidence and mortality of nasopharyngeal carcinoma. *Cancer Lett.* 2016;374(1): 22-30.
2. Jia WH, Feng BJ, Xu ZL, et al. Familial risk and clustering of nasopharyngeal carcinoma in Guangdong, China. *Cancer.* 2004;101(2): 363-369.
3. Liu Z, Chang ET, Liu Q, et al. Quantification of familial risk of nasopharyngeal carcinoma in a high-incidence area. *Cancer.* 2017;123(14): 2716-2725.
4. Dai J, Shen W, Wen W, et al. Estimation of heritability for nine common cancers using data from genome-wide association studies in Chinese population. *Int J Cancer.* 2017;140(2):329-336.
5. Yu G, Hsu WL, Coghill AE, et al. Whole-exome sequencing of nasopharyngeal carcinoma families reveals novel variants potentially involved in nasopharyngeal carcinoma. *Sci Rep.* 2019;9(1):9916.
6. Welch PL, Owens KN, King MC. Insights into the functions of BRCA1 and BRCA2. *Trends Genet.* 2000;16(2):69-74.
7. Mahdavi M, Nassiri M, Kooshyar MM, et al. Hereditary breast cancer; genetic penetrance and current status with BRCA. *J Cell Physiol.* 2019; 234(5):5741-5750.
8. Lu SJ, Day NE, Degos L, et al. Linkage of a nasopharyngeal carcinoma susceptibility locus to the HLA region. *Nature.* 1990;346(6283): 470-471.
9. Su WH, Hildesheim A, Chang YS. Human leukocyte antigens and Epstein-Barr virus-associated nasopharyngeal carcinoma: old associations offer new clues into the role of immunity in infection-associated cancers. *Front Oncol.* 2013;3:299.
10. Ramayanti O, Juwana H, Verkuijlen SA, et al. Epstein-Barr virus mRNA profiles and viral DNA methylation status in nasopharyngeal brushings from nasopharyngeal carcinoma patients reflect tumor origin. *Int J Cancer.* 2017;140(1):149-162.
11. Tsao SW, Tsang CM, Lo KW. Epstein-Barr virus infection and nasopharyngeal carcinoma. *Philos Trans R Soc Lond B Biol Sci.* 2017; 372(1732).



12. Chien YC, Chen JY, Liu MY, et al. Serologic markers of Epstein-Barr virus infection and nasopharyngeal carcinoma in Taiwanese men. *New Engl J Med*. 2001;345(26):1877-1882.
13. Zeng Y, Zhang LG, Wu YC, et al. Prospective studies on nasopharyngeal carcinoma in Epstein-Barr virus IgA/VCA antibody-positive persons in Wuzhou City, China. *Int J Cancer*. 1985;36(5):545-547.
14. Chen Y, Zhao W, Lin L, et al. Nasopharyngeal Epstein-Barr virus load: an efficient supplementary method for population-based nasopharyngeal carcinoma screening. *PLoS One*. 2015;10(7):e0132669.
15. Zheng XH, Lu LX, Li XZ, Jia WH. Quantification of Epstein-Barr virus DNA load in nasopharyngeal brushing samples in the diagnosis of nasopharyngeal carcinoma in southern China. *Cancer Sci*. 2015;106(9):1196-1201.
16. He YQ, Liao XY, Xue WQ, et al. Association between environmental factors and oral Epstein-Barr virus dna loads: a multicenter cross-sectional study in China. *J Infect Dis*. 2019;219(3):400-409.
17. Xue WQ, He YQ, Liao XY, et al. Decreased oral Epstein-Barr virus DNA loads in patients with nasopharyngeal carcinoma in southern China: a case-control and a family-based study. *Cancer Med*. 2018;7:3453-3464.
18. Lo YM, Chan LY, Lo KW, et al. Quantitative analysis of cell-free Epstein-Barr virus DNA in plasma of patients with nasopharyngeal carcinoma. *Cancer Res*. 1999;59(6):1188-1191.
19. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44(7):821-824.
20. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet*. 1998;62(5):1198-1211.
21. Qin HD, Jia WH, Zhang LL, et al. Elevated Epstein-Barr virus seroreactivity among unaffected members of families with nasopharyngeal carcinoma. *J Med Virol*. 2011;83(10):1792-1798.
22. Mathias RA, Deepa M, Deepa R, Wilson AF, Mohan V. Heritability of quantitative traits associated with type 2 diabetes mellitus in large multiplex families from South India. *Metab: Clin Exp*. 2009;58(10):1439-1455.
23. Souto JC, Pena G, Ziyatdinov A, et al. A genomewide study of body mass index and its genetic correlation with thromboembolic risk. Results from the GAIT project. *Thromb Haemost*. 2014;112(5):1036-1043.
24. Kim Y, Lee Y, Lee S, et al. On the estimation of heritability with family-based and population-based samples. *BioMed Res Int*. 2015;2015:671349.
25. Stevens SJ, Verkuijlen SA, Hariwiyanto B, et al. Noninvasive diagnosis of nasopharyngeal carcinoma: nasopharyngeal brushings reveal high Epstein-Barr virus DNA load and carcinoma-specific viral BARF1 mRNA. *Int J Cancer*. 2006;119(3):608-614.
26. Zheng X-H, Wang R-Z, Li X-Z, et al. Detection of methylation status of Epstein-Barr virus DNA C promoter in the diagnosis of nasopharyngeal carcinoma. *Cancer Sci*. 2020;111(2):592-600.
27. Song Y, Schwager MJ, Backer V, et al. Environment changes genetic effects on respiratory conditions and allergic phenotypes. *Sci Rep*. 2017;7(1):6342.
28. Renner B, Mueller CA, Shephard A. Environmental and non-infectious factors in the aetiology of pharyngitis (sore throat). *Inflammation Res*. 2012;61(10):1041-1052.
29. Graham JP, Corella Barud V, Avitia Diaz R, Gurian P. The in-home environment and household health: a cross-sectional study of informal urban settlements in northern Mexico. *Int J Env Res Public Health*. 2005;2(3-4):394-402.
30. Besson C, Amiel C, Le-Pendeven C, et al. Strong correlations of anti-viral capsid antigen antibody levels in first-degree relatives from families with Epstein-Barr virus-related lymphomas. *J Infect Dis*. 2009;199(8):1121-1127.
31. Tse KP, Su WH, Chang KP, et al. Genome-wide association study reveals multiple nasopharyngeal carcinoma-associated loci within the HLA region at chromosome 6p21.3. *Am J Hum Genet*. 2009;85(2):194-203.
32. Bei JX, Li Y, Jia WH, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet*. 2010;42(7):599-603.
33. Ruan HL, Qin HD, Shugart YY, et al. Developing genetic epidemiological models to predict risk for nasopharyngeal carcinoma in high-risk population of China. *PLoS One*. 2013;8(2):e56128.
34. Bei J-X, Su W-H, Ng C-C, et al. A GWAS meta-analysis and replication study identifies a novel locus within CLPTM1L/TERT associated with nasopharyngeal carcinoma in individuals of Chinese ancestry. *Cancer Epidemiol Biomark Prev*. 2016;25(1):188-192.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Zheng M-Q, Wang T-M, Liao Y, et al. Nasopharyngeal Epstein-Barr virus DNA loads in high-risk nasopharyngeal carcinoma families: Familial aggregation and host heritability. *J Med Virol*. 2020;92:3717-3725. <https://doi.org/10.1002/jmv.26198>